

대한민국 특허청

KOREAN INTELLECTUAL
PROPERTY OFFICE

별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

CERTIFIED COPY OF
PRIORITY DOCUMENT

출원번호 : 10-2003-0020059
Application Number

출원년월일 : 2003년 03월 31일
Date of Application MAR 31, 2003

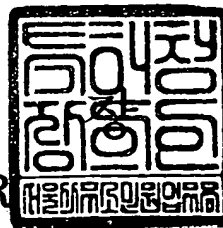
출원인 : 삼성전자주식회사
Applicant(s) SAMSUNG ELECTRONICS CO., LTD.



2003 년 10 월 16 일

특허청

COMMISSIONER



BEST AVAILABLE COPY



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0007
【제출일자】	2003.03.31
【국제특허분류】	G10L
【발명의 명칭】	두개의 센서를 이용한 목적원별 신호 분리방법 및 장치
【발명의 영문명칭】	Method and apparatus for blind source separation using two sensors
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2003-003435-0
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2003-003436-7
【발명자】	
【성명의 국문표기】	최창규
【성명의 영문표기】	CHOI, Chang Kyu
【주민등록번호】	700309-1002415
【우편번호】	110-011
【주소】	서울특별시 종로구 구기동 62-1
【국적】	KR
【심사청구】	청구
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인 이영필 (인) 대리인 이해영 (인)



1020030020059

출력 일자: 2003/10/21

【수수료】

【기본출원료】 20 면 29,000 원

【가산출원료】 22 면 22,000 원

【우선권주장료】 0 건 0 원

【심사청구료】 16 항 621,000 원

【합계】 672,000 원

【첨부서류】 1. 요약서·명세서(도면)_1통

【요약서】

【요약】

두 개의 센서로부터 수신되어 주파수영역으로 변환된 제1 및 제2 센서수신신호로부터 적어도 두 개 이상의 목적원별로 신호를 분리하는 방법 및 장치가 개시된다. 그 방법은 (a) 제1 및 제2 센서수신신호 중 적어도 하나에 대하여 프레임별로 광역 신호부재확률과 해당 프레임의 각 주파수대역에 대한 국지 신호부재확률을 계산하는 단계, (b) (a) 단계에서 계산된 국지 신호부재확률에 의해 신호가 존재한다고 판단되는 주파수대역의 신호값에 대하여 생성되는 감쇠와 지연시간을 이용하여 각 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 단계, (c) (a) 단계에서 계산된 광역 신호부재확률에 의해 각 주파수대역별로 잡음신호를 제거한 스펙트럼벡터를 추정하는 단계, 및 (d) (b) 단계에서 얻어진 목적원별 레이블벡터와 (c) 단계에서 얻어진 각 주파수대역에 대하여 추정된 스펙트럼벡터를 각각 승산하여 목적원별로 분리된 신호를 획득하는 단계를 포함한다.

【대표도】

도 1

【명세서】**【발명의 명칭】**

두개의 센서를 이용한 목적원별 신호 분리방법 및 장치(Method and apparatus for blind source separation using two sensors)

【도면의 간단한 설명】

도 1은 본 발명에 따른 목적원별 신호 분리장치의 제1 실시예의 구성을 보여주는 블록도,

도 2는 도 1에 있어서 신호부재확률 계산부의 동작을 설명하는 흐름도,

도 3은 도 1에 있어서 목적원 식별부의 동작을 설명하는 흐름도,

도 4a 및 도 4b는 도 3에 도시된 목적원 식별부에서 클러스터링한 결과를 보여주는 도면

,

도 5는 도 1에 있어서 신호 추정부의 동작을 설명하는 흐름도,

도 6은 도 1에 있어서 목적원 분리부의 동작을 설명하는 흐름도,

도 7은 본 발명에 따른 목적원별 신호 분리장치의 제2 실시예의 구성을 보여주는 블록도

,

도 8은 본 발명에 따른 목적원별 신호 분리 알고리즘의 성능을 평가하기 위한 실험환경을 나타낸 도면, 및

도 9a 내지 도 9c는 본 발명에 따른 목적원별 신호 분리 알고리즘의 성능을 평가하기 위한 목적원별 신호, 마이크로폰 수신신호 및 분리된 목적원별 신호를 나타내는 파형도이다.

【발명의 상세한 설명】**【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

- <10> 본 발명은 신호원 분리(Blind Source Separation)에 관한 것으로서, 특히 두 개의 마이크로폰 수신신호로부터 잡음을 실시간으로 제거함과 동시에 임의의 N개의 목적원별로 신호를 분리하기 위한 방법 및 장치에 관한 것이다.
- <11> 최근 들어 건강, 안전, 홈 네트워크, 엔터테인먼트 등의 필요성에 의해 더욱 더 사람들의 관심을 끌고 있는 이동 로봇을 동작시키기 위해서는 사람과 로봇 간의 상호작용(Human-Robot Interaction)을 필요로 한다. 즉, 이동 로봇은 사람이 하는 것과 마찬가지로 비전시스템을 가지고 사람을 인식하고 주변 상황을 인지할 수 있어야 하며, 자신의 주위에서 말하는 사람의 위치를 알아야 하고 그 사람이 하는 말을 이해할 수 있어야 한다.
- <12> 이동 로봇에 있어서 음성 입력시스템은 사람과 로봇 간의 상호작용에 필수적일 뿐만 아니라 자율 주행에도 중요한 이슈가 된다. 실내 환경에서 이동 로봇의 음성 입력시스템에 야기되는 중요한 문제는 잡음, 반향 및 거리이다. 실내 환경에서는 여러가지 잡음원과 벽이나 기타 사물에 의한 반향이 존재한다. 거리에 따라서 음성의 저주파 성분은 고주파 성분에 비하여 더 많이 감쇄되는 특징이 있다. 따라서, 가정의 실내 환경에서 사람과 로봇 간의 상호작용에 필요한 음성 입력시스템은 자율 주행 이동 로봇이 수 미터 떨어진 거리에서 사용자의 평상적인 목소리를 받아들여 음성인식에 직접적으로 사용될 수 있어야 한다.
- <13> 이러한 음성 입력시스템에서는 음질 및 음성인식율을 향상시키기 위하여 적어도 두개 이상의 마이크로폰으로 이루어진 마이크로폰 어레이를 널리 사용하게 되고, 마이크로폰 어레이로

부터 입력되는 음성신호에 포함된 잡음성분을 제거하기 위하여 단일채널 음성강조기법(Single channel speech Enhancement), 적응적 음향잡음 제거기법(Adaptive acoustic Noise Canceling Method), 일반화된 사이드로브 제거기법(Generalized Sidelobe Canceling Method) 또는 블라인드신호 분리기법(Blind Signal Separatuion) 등과 같은 방법들을 사용하고 있다.

<14> 단일채널 음성강조기법은 하나의 마이크로폰을 사용하며 정적인 배경잡음(stationary background noise)과 같이 잡음의 통계적 특성이 시간적으로 변하지 않은 경우에만 그 성능을 발휘한다. B. Widrow et al.의 『"Adaptive Noise Canceling: Principles and Applications," *Proceedings of IEEE*, vol. 63, no. 12, pp. 1692-1716, 1975』에 개시된 음향잡음 제거기법은 두개의 마이크로폰을 사용하는데, 그 중 하나는 잡음만을 수음하는 참조 마이크로폰이므로, 잡음만을 수음할 수 없는 경우 또는 참조 마이크로폰으로 수음한 잡음 이외의 잡음이 혼입되는 경우에는 그 성능이 급격히 저하된다. O. Hoshuyama et al.의 『"A Robust Adaptive Beamformer For Microphone Arrays With A Blocking Matrix Using Constrained Adaptive Filters," in *IEEE Trans. Signal Processing*, vol. 47, no. 10, pp. 2677-2684, 1999』에 개시된 일반화된 사이드로브 제거기법은 음성 활성화 검출기(Voice Activity Detector)를 필요로 할 뿐 아니라 목적원신호가 잡음신호와 함께 제거되는 단점이 있다.

<15> 한편, 블라인드신호 분리기법의 일종인 DUET(Degenerate Unmixing Estimation Technique)와 관련된 종래기술은 A. Jourjine, S. Rickard 및 O. Yilmaz에 의한 『"Blind separation of disjoint orthogonal signals: Demixing n sources from 2 mixtures," in *Proc. Int. Conf. on Acoust., Speech, Signal Processing*, 2000. vol. 5, pp. 2985-2988』, S. Rickard, R. Balan 및 J. Rosca에 의한 『"Real-time time-frequency based blind source separation," in

Proc. Int. Conf. on Independent Component Analysis and Blind Signal Separation, 2001, pp. 651-656』, S. Rickard 및 O. Yilmaz에 의한 『"On the approximate w-disjoint orthogonality of speech," in *Proc. ICASSP 2002*, pp. 529-532』 등을 들 수 있다. 이러한 종래의 DUET는 두 음성신호 $s_1(t)$ 와 $s_2(t)$ 의 주파수성분은 서로 겹치지 않는다는, 즉 하나의 주파수대역에는 단 하나의 음성신호 성분만 존재한다는 w-디스조인트 직교성이라는 기본 가정을 이용한다. 그러나, 혼합신호에 잡음이 섞인 경우 및 w-디스조인트 직교성 가정이 성립하지 않는 경우, 즉 전 주파수대역에 걸친 스펙트럼을 갖는 백색 가우시안 잡음이나 비교적 넓은 주파수대역에 걸쳐 스펙트럼이 존재하는 팬 잡음이 음성신호와 혼합된 경우에는 분리성능이 급격히 저하되는 단점이 있다.

【발명이 이루고자 하는 기술적 과제】

- <16> 따라서, 본 발명이 이루고자 하는 기술적 과제는 두 개의 마이크로폰 수신신호로부터 잡음을 실시간으로 제거함과 동시에 임의의 N개의 목적원별로 신호를 분리하기 위한 방법을 제공하는데 있다.
- <17> 본 발명이 이루고자 하는 다른 기술적 과제는 상기 목적원별 신호 분리방법을 실현하는데 가장 적합한 장치를 제공하는데 있다.
- <18> 상기 기술적 과제를 달성하기 위하여 본 발명에 따른 두 개의 센서를 이용한 목적원별 신호 분리방법은 두 개의 센서로부터 수신되어 주파수영역으로 변환된 제1 및 제2 센서수신신호로부터 적어도 두 개 이상의 목적원별로 신호를 분리하는 방법에 있어서, (a) 상기 제1 및 제2 센서수신신호 중 적어도 하나에 대하여 프레임별로 광역 신호부재확률과 해당 프레임의 각 주파수대역에 대한 국지 신호부재확률을 계산하는 단계; (b) 상기 (a) 단계에서 계산된 국지 신호부재확률에 의해 신호가 존재한다고 판단되는 주파수대역의 신호값에 대하여 생성되는 감

식와 지연시간을 이용하여 각 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 단계; (c) 상기 (a) 단계에서 계산된 광역 신호부재확률에 의해 각 주파수대역별로 잡음신호를 제거한 스펙트럼벡터를 추정하는 단계; 및 (d) 상기 (b) 단계에서 얻어진 목적원별 레이블벡터와 상기 (c) 단계에서 얻어진 각 주파수대역에 대하여 추정된 스펙트럼벡터를 각각 승산하여 목적원별로 분리된 신호를 획득하는 단계를 포함한다.

<19> 상기 다른 기술적 과제를 달성하기 위하여 본 발명은 두 개의 센서로부터 수신되어 주파수영역으로 변환된 제1 및 제2 센서수신신호로부터 적어도 두 개 이상의 목적원별로 신호를 분리하는 장치에 있어서, 상기 제1 및 제2 센서수신신호 중 적어도 하나에 대하여 프레임별로 광역 신호부재확률과 해당 프레임의 각 주파수대역에 대한 국지 신호부재확률을 계산하는 신호부재확률 계산부; 상기 신호부재확률 계산부에서 계산된 광역 신호부재확률에 의해 각 주파수대역별로 잡음신호를 제거한 스펙트럼벡터를 추정하는 신호추정부; 상기 신호부재확률 계산부에서 계산된 국지 신호부재확률에 의해 신호가 존재한다고 판단되는 주파수대역의 신호값에 대하여 생성되는 감쇠와 지연시간을 이용하여 각 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 목적원 식별부; 및 상기 목적원 식별부에서 얻어진 목적원별 레이블벡터와 상기 신호추정부에서 얻어진 각 주파수대역에 대하여 추정된 스펙트럼벡터를 각각 승산하여 목적원별로 분리된 신호를 목적원별 신호분리부를 포함한다.

【발명의 구성 및 작용】

<20> 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 일실시예에 대하여 상세히 설명하기로 한다.

<21> 본 발명에 따른 목적원별 신호 분리방법 및 장치에 대하여 설명하기에 앞서 본 발명에 적용된 원칙들을 살펴보기로 한다.

<22> 먼저, 작은 반향이 존재하는 환경에서 목적원신호($s_j(t)$)와 마이크로폰 수신신호($x_j(t)$)는 다음 수학적 식 1과 같이 나타낼 수 있다.

<23>

$$x_1(t) = \sum_{j=1}^N s_j(t) + n_1(t)$$

【수학적 식 1】

<24>

$$x_2(t) = \sum_{j=1}^N a_j s_j(t - d_j) + n_2(t)$$

【수학적 식 2】

<25> 상기 수학적 식 1 및 2를 참조하면, $n_1(t)$ 는 배경 잡음원으로부터 제1 마이크로폰 수신신호($x_1(t)$)에 추가되는 잡음, $n_2(t)$ 는 다른 전파경로를 통해 제2 마이크로폰 수신신호($x_2(t)$)에 추가되는 잡음, a_j 는 크기 감쇠비, d_j 는 도착방향(Direction of Arrival)으로 인한 마이크로폰들 간의 지연시간을 각각 나타낸다.

<26> 이와 같은 시간영역의 신호에 대하여 푸리에 변환, 예를 들면 단구간 푸리에변환(Short-Term Fourier Transform)을 수행하여 주파수영역의 신호로 변환한다. 주파수영역에 있어서, 제1 마이크로폰 수신신호($X(m)$)가 목적원신호($S(m)$)와 부가잡음($N(m)$)의 합으로 표현된다고 가정하면, 광역가설(H_0 , H_1)을 도입한 통계모델을 도출할 수 있다. 여기서, H_0 및 H_1 는 각각 m 번째 프레임에서 목적원신호의 부재 및 존재를 각각 나타내며, 다음 수학적 식 3과 같이 나타낼 수 있다.

<27>

$$\begin{aligned} H_{0,k} : & X_k(m) = N_k(m) \\ H_{1,k} : & X_k(m) = S_k(m) + N_k(m) \end{aligned}$$

【수학적 식 3】

<28> H_1 에 존재하는 목적원신호($S(m)$)는 개개 목적원신호들의 합으로서 다음 수학적 식 4와 같이 나타낼 수 있다.

<29>

$$\text{【수학적 식 4】 } S(m) = \sum_{j=1}^N S_j(m)$$

<30> 또한, 목적원신호의 부재 및 존재는 각 주파수빈에 대하여 발생하기 때문에

국지가설($H_{0,k}$, $H_{1,k}$)을 도입한 통계모델을 도출할 수 있다. 여기서, $H_{0,k}$ 및 $H_{1,k}$ 는 각각 m 번째 프레임의 k 번째 주파수빈에서 목적원 신호성분의 부재 및 존재를 각각 나타내며, 다음 수학적 식 5와 같이 나타낼 수 있다.

<31>

$$\begin{aligned} H_0 : & \mathbf{X}(m) = \mathbf{N}(m) \\ \text{【수학적 식 5】 } H_1 : & \mathbf{X}(m) = \mathbf{S}(m) + \mathbf{N}(m) \end{aligned}$$

<32> $H_{1,k}$ 에 존재하는 목적원신호($S_k(m)$)는 개개 소스신호 성분들의 합으로서 다음 수학적 식 6과 같이 나타낼 수 있다.

<33> 【수학적 식 6】 $S_k(m) = \sum_{j=1}^N S_k^{(j)}(m)$

<34> 여기서 $S_k^{(j)}(m)$ 는 m 번째 프레임의 k 번째 주파수빈, 즉 주파수대역에 존재하는 j 번째 목적원 신호성분을 나타낸다. w -디스조인트 직교성(w -disjoint orthogonality) 조건을 만족한다고 가정하면, k 번째 주파수대역은 하나의 음원에 대해서만 점유되므로 임의의 j 에 대하여 $S_k(m) = S_k^{(j)}(m)$ 으로 나타낼 수 있다. w -디스조인트 직교성에 대해서는 S. Rickard 와 O. Yilmaz에 의한 논문 『"On the approximate w -disjoint orthogonality of speech," in Proc. Int. Conf. on Acoust., Speech, Signal Processing, 2002, vol. 1, pp. 529-532』에 자세하게 개시되어 있다. 또한, $X_k(m)$ 과 $N_k(m)$ 은 평균이 제로(0)인 복소 가우시안 확률분포를 가진다고 가정한다.

<35> 상술한 바와 같은 기본원칙을 바탕으로 하여 본 발명에 대하여 세부적으로 설명하면 다음과 같다.



<36> 도 1은 본 발명에 따른 목적원별 신호 분리장치의 일실시예의 구성을 나타낸

블록도로서, 신호부재확률(Signal Absence Probability) 계산부(11), 목적원 식별부(13), 신호 추정부(15) 및 목적원별 신호분리부(17)로 이루어진다. 본 발명에 따른 목적원별 신호 분리장치는 M개의 마이크로폰 수신신호로부터 N개의 독립된 목적원을 분리하기 위한 것으로서, 이하에서는 설명의 편의를 위하여 M=2인 경우를 예를 들어 설명하기로 한다.

<37> 도 1을 참조하면, 신호부재확률 계산부(11)에서는 두 개의 센서 예를 들면, 마이크로폰 으로부터 수신되어 주파수영역으로 변환된 제1 및 제2 마이크로폰 수신신호 중 하나에 대하여 m 번째 프레임에 대한 광역 신호부재확률($p(H_0 | X(m))$)과 m 번째 프레임의 각 주파수빈에 대한 국지 신호부재확률($p(H_{0,k} | X_k(m))$)을 계산한다. 이에 대하여 도 2를 참조하여 좀 더 세부적으로 설명하면 다음과 같다.

<38> 도 2에 있어서, 21 단계에서는 소정 프레임의 (L/2)개의 주파수빈에 대하여 가능비 (likelihood ratio, $\Lambda_k(m)$)를 다음 수학식 7에서와 같이 계산한다. 여기서, L은 푸리에 변환의 포인트 수를 나타낸다.

<39>
$$\Lambda_k(m) = \frac{p(X_k(m) | H_{1,k})}{p(X_k(m) | H_{0,k})}$$

【수학식 7】

<40> 여기서,

<41>
$$p(X_k(m) | H_{1,k}) = \frac{1}{\pi(\sigma_{S_t}^2(m) + \sigma_{N_t}^2(m))} \exp \left[-\frac{|X_k(m)|^2}{(\sigma_{S_t}^2(m) + \sigma_{N_t}^2(m))} \right]$$

<42>
$$p(X_k(m) | H_{0,k}) = \frac{1}{\pi\sigma_{N_t}^2(m)} \exp \left[-\frac{|X_k(m)|^2}{\sigma_{N_t}^2(m)} \right]$$

<43> 이다. $p(X_k(m) | H_{1,k})$ 는 국지 신호존재 가설에서 m 번째 프레임의 k 번째 주파수빈의 확률값이고, $p(X_k(m) | H_{0,k})$ 는 국지 신호부재 가설에서 m 번째 프레임의 k 번째 주파수빈의 확률값으로서, 목적원신호전력($\sigma^2_{s_k}$)과 잡음신호전력($\sigma^2_{n_k}$)으로부터 구해질 수 있다.

<44> 23 단계에서는 상기 21 단계에서 구한 $(L/2)$ 개의 가능비를 각각 다음 수학적식 8에서와 같이 표현되는 소정의 어프라이오리 확률(a priori probability, q_k)과 승산한다.

<45> **【수학적식 8】**
$$q_k = \frac{p(H_{1,k})}{p(H_{0,k})}$$

<46> 여기서, $p(H_{1,k})$ 는 k 번째 주파수빈에서 잡음과 목적원 신호가 공존할 확률, $p(H_{0,k})$ 는 k 번째 주파수빈에서 잡음신호만 존재할 확률을 각각 나타낸다.

<47> 25 단계에서는 상기 23 단계에서 구한 승산결과 각각을 소정값, 예를 들면 '1'과 가산한다.

<48> 27 단계에서는 상기 25 단계에서 구한 가산결과 각각에 대하여 역수를 취하여 다음 수학적식 9와 같이 표현되는 m 번째 프레임의 $(L/2)$ 개의 주파수빈에 대한 국지 신호부재확률 ($p(H_{0,k} | X_k(m))$)을 계산한다.

<49> **【수학적식 9】**
$$p(H_{0,k} | X_k(m)) = \frac{1}{1 + q_k \Lambda_k(m)}$$

<50> 29 단계에서는 상기 27 단계에서 구한 $(L/2)$ 개의 국지 신호부재확률 ($p(H_{0,k} | X_k(m))$)을 승산하여 다음 수학적식 10과 같이 표현되는 m 번째 프레임에 대한 광역 신호부재확률 ($p(H_0 | X(m))$)을 계산한다.

<51>

$$\begin{aligned}
 p(H_0|\mathbf{X}(m)) &= \frac{p(H_0, \mathbf{X}(m))}{p(\mathbf{X}(m))} \\
 &= \frac{1}{\prod_{k=1}^M [1 + q_k \Lambda_k(m)]}
 \end{aligned}$$

【수학식 10】

<52>

한편, 상기 수학식 10을 도출하기 위하여 다음 수학식 11 및 12가 사용되어진다.

<53>

$$p(H_0, \mathbf{X}(m)) = \prod_{k=1}^M [p(X_k(m)|H_{0,k})p(H_{0,k})]$$

【수학식 11】

<54>

$$\begin{aligned}
 p(\mathbf{X}(m)) &= \prod_{k=1}^M p(X_k(m)) \\
 &= \prod_{k=1}^M [p(X_k(m)|H_{0,k})p(H_{0,k}) \\
 &\quad + p(X_k(m)|H_{1,k})p(H_{1,k})]
 \end{aligned}$$

【수학식 12】

<55>

이때, 단구간 푸리에변환 계수들은 서로 상관되어 있지 않기 때문에 각 주파수빈에서의 스펙트럼 성분 즉, $X_k(m)$ 은 통계적으로 독립적이라고 가정한다.

<56>

다시 도 1로 돌아가서, 목적원 식별부(13)에서는 신호부재확률 계산부(11)에서 계산된 제1 또는 제2 마이크로폰 수신신호의 프레임별 국지 신호부재확률을 입력으로 하여 미리 정의된 제1 문턱치와 비교하고, 비교 결과에 따라서 각 프레임의 해당 주파수빈 즉, 주파수대역에 목적원 신호가 존재하는지 여부를 판단한다. 이후, 목적원 신호가 존재한다고 판단되는 주파수대역의 신호값에 대하여 크기 감쇠와 지연시간을 생성하고, 크기 감쇠와 지연시간으로 이루어지는 혼합 파라미터를 이용하여 목적원의 수 및 각 목적원에 속하는 주파수대역을 구분한다. 목적원 식별부(13)는 프레임 단위로 상기한 목적원 식별과정을 수행한다. 이를 도 3을 참조하여 좀 더 세부적으로 설명하면 다음과 같다.

<57> 도 3에 있어서, 31 단계에서는 m 번째 프레임의 국지 신호부재확률을 제1 문턱치와 비교하여 목적원신호가 존재하는 주파수대역을 판별한다. 즉, 목적원이 w -디스조인트 직교성을 만족하는 것으로 가정하면 N 개의 목적원 중에서 단지 하나의 목적원 신호성분이 k 번째 주파수대역에 존재한다. 따라서, 상기 수학식 9에 의해 계산된 국지 신호부재확률을 제1 문턱치와 비교함으로써 어떤 주파수대역이 목적원신호를 가지고 있는지를 판단할 수 있다. 예를 들어, 국지 신호부재확률이 제1 문턱치보다 크거나 같으면 해당 주파수대역에 목적원신호가 존재하지 않는 것으로 판단하고, 제1 문턱치보다 작으면 해당 주파수대역에 목적원신호가 존재하는 것으로 판단한다. 이와 같이 목적원신호가 존재한다고 판단된 주파수대역은 임의의 목적원 j 에 대하여 다음 수학식 13이 성립된다.

<58> **【수학식 13】**
$$\begin{bmatrix} X_{1,k}(m) \\ X_{2,k}(m) \end{bmatrix} \approx \begin{bmatrix} 1 \\ a_j e^{-i\omega d_j} \end{bmatrix} S_k^{(j)}(m)$$

<59> 여기서, ω 는 $(2\pi k)/L$ 이며, L 은 푸리에변환의 포인트의 수를 나타낸다.

<60> 32 단계에서는 상기 31 단계에 의해 목적원신호가 존재하는 것으로 판단된 각 주파수대역에 대하여 제1 마이크로폰 수신신호($X_{1,k}(m)$)와 제2 마이크로폰 수신신호($X_{2,k}(m)$)의 비를 분석하여 각 주파수대역에 대하여 감쇠(a_j) 및 지연시간(d_j)을 계산한다. 이와 같이 크기 감쇠와 지연시간으로 이루어지는 k 번째 주파수대역의 혼합 파라미터($\mu^{(k)}$)는 다음 수학식 14와 같이 표현할 수 있다.

<61> **【수학식 14】**
$$\begin{aligned} \mu^{(k)} &\equiv (a(k, m), d(k, m)) \\ &= \left(\left| \frac{X_{2,k}(m)}{X_{1,k}(m)} \right|, -\frac{1}{\omega} \angle \frac{X_{2,k}(m)}{X_{1,k}(m)} \right) \end{aligned}$$

- <62> 상기 수학적 식 14를 이용하면 m 번째 프레임에서 목적원신호가 존재하는 각 주파수대역들, 예를 들면 P 개의 혼합 파라미터를 구할 수 있다.
- <63> 33 단계에서는 소프트 K-평균(K-means) 알고리즘을 이용하여 상기 32 단계에서 구한 P 개의 혼합 파라미터를 클러스터링하기 위하여 먼저 현재 프레임이 첫 번째 프레임인가를 판단한다. 34 단계에서는 상기 33 단계에서의 판단결과, 현재 프레임이 첫 번째 프레임인 경우에는 K 개 클러스터의 중심(μ_j)을 랜덤 값으로 초기화시킨다. 35 단계에서는 상기 33 단계에서의 판단결과, 현재 프레임이 첫 번째 프레임이 아닌 경우에는 클러스터의 중심 즉, 평균값을 이전 프레임의 클러스터 중심으로 초기화시킨다. 한편, 클러스터의 수(K)는 2 내지 N 으로 설정한 다음 소정의 클러스터링 알고리즘을 수행하고, 각각의 K 에 대하여 David L. Davies 및 Donald W. Bouldin에 의한 『"A Cluster Separation Measure," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-1, no. 2, April 1979』의 정의 5에 제시된 R 값을 계산하고, R 값이 가장 작은 K 가 클러스터의 수에 해당한다.
- <64> 36 단계에서는 혼합 파라미터($\mu^{(k)}$, 여기서 k 는 1 내지 P)와 각 클러스터의 중심(μ_j)을 이용하여 각 주파수대역이 각 클러스터(j)에 속할 소속도($r_j^{(k)}$)를 다음 수학적 식 15에서와 같이 계산하고, 가장 큰 값을 갖는 클러스터에 해당 주파수대역을 소속시킨다.
- <65>
$$r_j^{(k)} = \frac{\exp(-\beta d(\mu_j, \mu^{(k)}))}{\sum \exp(-\beta d(\mu_j, \mu^{(k)}))}$$
【수학적 식 15】
- <66> 여기서, $d(\mu_j, \mu^{(k)})$ 는 클러스터의 중심과 혼합 파라미터간의 거리 연산자, β 는 감쇠 파라미터로서 잡음분산에 반비례하는 값이다.

<67> 37 단계에서는 상기 수학식 15에서 계산된 소속도에 따라 클러스터가 임의의 주파수대역이 소속되면 각 클러스터(j)의 중심(μ_j)을 다음 수학식 16에서와 같이 갱신한다.

<68>

$$\mu_j = \frac{\sum_k r_j^{(k)} \mu^{(k)}}{\sum_k r_j^{(k)}} \equiv (a_j(m), d_j(m))$$

【수학식 16】

<69> 여기서, $a_j(m)$ 및 $d_j(m)$ 은 각각 m 번째 프레임에서 j 번째 목적원에 대한 감쇠 및 지연시간을 나타낸다.

<70> 38 단계에서는 수렴 여부를 판단하는데, 상기 수학식 15 및 16에 의한 소속도 계산 및 클러스터 중심 갱신과정을 소속도의 변화가 없을 때까지 또는 정해진 반복횟수만큼 수행한다. 즉 상기 38 단계에서의 판단결과, 수렴이 되지 않은 경우 상기 36 단계로 복귀하여 상기 수학식 16에 의해 갱신된 클러스터의 중심을 이용하여 재차 소속도 계산 및 클러스터 중심 갱신과정을 반복 수행한다.

<71> 39 단계에서는 상기 38 단계에서의 판단결과, 수렴이 이루어진 경우 클러스터의 수를 목적원의 수로 간주하고, 목적원별로 레이블벡터를 생성한다.

<72> 도 4a 및 도 4b는 목적원 식별부(13)에서 클러스터링한 결과를 각각 2 차원 및 3 차원 a-d 공간상에 표시한 것으로서, 여기서는 3개의 클러스터가 생성되었으므로 3개의 목적원이 존재함을 알 수 있다.

<73> 한편, 이동평균(moving average)을 이용하여 최종적으로 현재 프레임의 클러스터 중심을 확정하는 방법은 목적원이 천천히 움직이는 경우에도 목적원을 잘 추종할 수 있다. 다음 프레임에서의 클러스터의 중심값은 현재 프레임에서 얻어진 중심값으로 초기화됨으로써 클러스터링 알고리즘의 수렴 속도를 증가시킬 수 있다.

- <74> P개의 혼합 파라미터를 클러스터링하기 위한 다른 실시예로서 하드 K-평균 알고리즘을 사용할 수도 있다. 하드 K-평균 알고리즘에서 소속도($r_j^{(k)}$)는 0 또는 1로만 나타내고, 그 결과 모든 데이터 포인트에 대하여 지수 함수를 구하지 않아도 되므로 계산 부담을 줄일 수 있다. 그러나, 소속도를 0 또는 1로만 나타내더라도 수렴 이후 클러스터의 중심값 계산시에는 각 주파수대역의 스펙트럼 성분값의 크기를 소속도로 대치하여 계산한다. 왜냐 하면, 동일한 크기의 잡음이 섞여도 스펙트럼 성분값이 큰 주파수대역은 전체 크기에 비해 섞이는 잡음의 크기가 상대적으로 작고, 스펙트럼 성분값이 작은 주파수대역은 전체 크기에 비해 섞이는 잡음의 크기가 상대적으로 크다. 따라서 잡음환경에서는 스펙트럼 성분값의 크기를 가중치로 두어 계산하는 것이 잡음에 더 둔감해질 수 있다.
- <75> 다시 도 1로 돌아가서, 신호추정부(15)에서는 신호부재확률 계산부(11)에서 계산된 광역 신호부재확률을 제2 문턱치와 비교하고, 광역 신호부재확률이 제2 문턱치보다 큰 프레임에 대해서는 잡음신호만 존재하는 것으로 판단하여 잡음신호 전력을 갱신하고, 모든 프레임에 대해서는 목적원신호 전력을 갱신함으로써 잡음신호가 제거된 목적원신호 성분을 발생시킨다. 이에 대하여 도 5를 참조하여 좀 더 세부적으로 설명하면 다음과 같다.
- <76> 도 5를 참조하면, 51 단계에서는 이전의 (m-1)번째 프레임에 대하여 후술하는 수학적식 25에서 계산된 제1 내지 제M 프라이오리 신호대잡음비를 이용하여 수학적식 26에서와 같이 m번째 프레임에 대한 예측 신호대잡음비를 계산한다.
- <77> 52 단계에서는 신호부재확률 계산부(11)에서 계산된 m번째 프레임에 대한 광역 신호부재확률을 제2 문턱치와 비교하여 현재 m번째 프레임에 잡음신호만 존재하는지 여부를 판단한다. 즉, m번째 프레임에 대한 광역 신호부재확률이 제2 문턱치보다 큰 경우 잡음신호만 존재하는 것으로 판단한다.



<78> 53 단계에서는 상기 52 단계에서의 판단결과, m 번째 프레임이 잡음신호만 존재하는 경우 다음 수학적식 17에서와 같이 잡음신호 전력($\sigma_{N_k}^2$)을 갱신한다.

<79> **【수학적식 17】**
$$\sigma_{N_k}^2(m) = \zeta_{N_k} \sigma_{N_k}^2(m-1) + (1 - \zeta_{N_k}) |N_k(m)|^2$$

<80> 상기 수학적식 17에 있어서 $0 < \zeta_{N_k} < 1$ 는 스무딩 파라미터이다.

<81> 54 단계에서는 상기 52 단계에서의 판단결과 잡음신호만 존재하여 53 단계에서 잡음신호 전력이 갱신된 잡음신호만 존재하는 프레임 또는 상기 52 단계에서의 판단결과 잡음신호와 목적원신호가 공존하는 프레임에 대하여 각 주파수대역에서의 스펙트럼 성분값을 추정한다. 이때, 잡음을 제거하기 위하여 최소평균자승에러(Minimum Mean Square Error) 추정 알고리즘을 이용한다.

<82> 즉, 신호부재확률 계산부(11)에서 제1 마이크로폰 수신신호에 대하여 신호부재확률을 계산한 경우, 제1 마이크로폰 수신신호($x_I(t)$)에서 m 번째 프레임의 k 번째 스펙트럼 성분을 $X_k(m)$ 이라 하면, $X_k(m)$ 은 다음 수학적식 18과 같이 나타낼 수 있다.

<83> **【수학적식 18】**
$$X_k(m) = S_k(m) + N_k(m)$$

<84> 여기서,

<85>
$$X_k(m) = R_k e^{-j\theta_k}$$

<86>
$$S_k(m) = A_k e^{-j\theta_k}$$

<87> 이다.

<88> 시공간에서의 신호의 파형 모양보다 각 주파수대역의 스펙트럼 진폭이 사람의 인지에 큰 영향을 미치는 것으로 알려져 있고, 이때 $a_k = \theta_k$ 가 최적의 위상값이다. 따라서, 최소평균자



승에러(Minimum Mean Square Error) 추정 알고리즘을 이용하면 X_k 로부터 A_k 를 추정하는 문제로 간략화될 수 있다. 이에 대해서는 Y. Ephraim 및 D. Malah에 의한 『"Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, no. 6, pp. 1109-1121, 1984』에 자세히 기재되어 있다.

<89> 각 주파수대역의 스펙트럼 성분값들은 서로 통계적으로 독립된 값들이므로 A_k 의 MMSE 추정값(\hat{A}_k)은 다음 수학적식 19 및 20에 의해 구해진다.

<90>

$$\begin{aligned}\hat{A}_k &= E\{A_k|\mathbf{X}\} \\ &= E\{A_k|X_k\} \\ &= p(H_{1,k}|X_k)E\{A_k|X_k, H_{1,k}\} \\ &= (1 - p(H_{0,k}|X_k))E\{A_k|X_k, H_{1,k}\} \\ &= \frac{q_k \Lambda_k}{1 + q_k \Lambda_k} E\{A_k|X_k, H_{1,k}\}\end{aligned}$$

【수학적식 19】

<91>

$$E\{A_k|X_k, H_{1,k}\} = \frac{\int_0^\infty \int_0^{2\pi} A_k p(X_k|A_k, \alpha_k) d\alpha_k dA_k}{\int_0^\infty \int_0^{2\pi} p(X_k|A_k, \alpha_k) d\alpha_k dA_k}$$

【수학적식 20】

<92> 상기 수학적식 20에 있어서 확률밀도함수 $p(X_k|A_k, a_k)$ 는 다음 수학적식 21과 같이 나타낼 수 있다.

<93>

$$p(X_k|A_k, \alpha_k) = \frac{1}{\pi \sigma_{N_k}^2} \exp\left(-\frac{1}{\sigma_{N_k}^2} |X_k - A_k e^{j\alpha_k}|^2\right)$$

【수학적식 21】

<94> 복소 가우시안 확률밀도함수는 마이크로폰 수신신호의 확률밀도함수와 잡음으로 간주되므로, $p(A_k)$ 는 레일레이(Rayleigh) 분포를 가지며, $p(a_k)$ 는 $p(A_k)$ 에 독립적이고 $[-\pi, \pi]$ 에 대하여 균일한 값이 된다. 따라서, $p(A_k, a_k)$ 는 다음 수학적식 22와 같이 나타낼 수 있다.

<95>

$$p(A_k, \alpha_k) = p(A_k)p(\alpha_k) = \frac{A_k}{\pi \sigma_{S_k}^2} \exp\left(-\frac{A_k^2}{\sigma_{S_k}^2}\right)$$

【수학식 22】

<96>

따라서, 상기 수학식 18에 표현된 A_k 의 MMSE 추정값(\hat{A}_k)은 다음 수학식 23에서와 같이 나타낼 수 있다.

<97>

$$\hat{A}_k = \frac{q_k \Lambda_k}{1 + q_k \Lambda_k} \cdot \Gamma(1.5) \frac{\sqrt{v_k}}{\gamma_k} \exp\left(-\frac{v_k}{2}\right) \cdot \left[(1 + v_k) I_0\left(\frac{v_k}{2}\right) + v_k I_1\left(\frac{v_k}{2}\right) \right] R_k$$

【수학식 23】

<98>

상기 수학식 23에 있어서, $\Gamma(\cdot)$ 은 $\Gamma(1.5) = \sqrt{\pi}/2$ 인 감마 함수이고, $I_0(\cdot)$ 및 $I_1(\cdot)$ 은 각각 0차와 1차의 수정된 베셀 함수를 나타내고, v_k 는 다음 수학식 24과 같이 정의된다.

<99>

$$v_k = \frac{\xi_k}{1 + \xi_k} (\gamma_k + 1)$$

【수학식 24】

<100>

또한, 상기 수학식 24에 포함된 프라이오리 신호대잡음비(ξ_k)와 예측 신호대잡음비(γ_k)는 각각 다음 수학식 25 및 26와 같이 정의된다.

<101>

$$\xi_k = \frac{\sigma_{S_k}^2}{\sigma_{N_k}^2}$$

【수학식 25】

<102>

$$\gamma_k = \frac{R_k^2}{\sigma_{N_k}^2} - 1$$

【수학식 26】

<103>

프라이오리 신호대잡음비가 매우 큰 경우 상기 수학식 23의 최소평균자승에러 추정값은 다음 수학식 27에서와 같이 위너(Wiener) 추정값과 유사하게 된다.

<104>

$$\hat{A}_k \approx \frac{\xi_k}{1 + \xi_k} R_k$$

【수학식 27】



<105> 55 단계에서는 상기 54 단계에서 계산된 각 주파수대역의 스펙트럼 성분(A_k)에 대한 MMSE 진폭추정값(\hat{A}_k)을 이용하여 다음 수학적식 28에서와 같이 목적원신호 전력($\sigma_{s_k}^2$)을 갱신한다.

<106> **【수학적식 28】** $\sigma_{s_k}^2 = |\hat{A}_k|^2$

<107> 즉, 실제 환경에서 목적원신호 및 잡음신호는 시간에 따라 끊임없이 변하게 된다. 이러한 변화를 본 발명에 따른 목적원 분리 알고리즘에 반영시키기 위해서 상기 수학적식 28 및 17에서와 같이 목적원신호 전력 및 잡음신호 전력을 매 프레임마다 갱신한다. 이와 같이 갱신된 잡음신호 전력($\sigma_{N_k}^2$)을 이용하여 상기 수학적식 26의 예측 신호대잡음비(γ_k)를 쉽게 갱신할 수 있다.

<108> 56 단계에서는 급격한 변화를 보이는 목적원신호에 대처하기 위하여 프라이오리 신호대 잡음비를 다음 수학적식 29에서와 같이 갱신된다.

<109> **【수학적식 30】**
$$\xi_k(m) = \zeta_{SNR} \frac{\hat{A}_k(m-1)}{\sigma_{N_k}^2(m-1)} + (1 - \zeta_{SNR}) P(\gamma_k(m))$$

<110> 상기 수학적식 30에 있어서 $\hat{A}_k(m-1)$ 은 (m-1) 번째 프레임에서 k 번째 스펙트럼 성분의 진폭 추정값이고, $P(\cdot)$ 는 다음 수학적식 31에서와 같이 정의되는 연산자이다.

<111> **【수학적식 31】**
$$P(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

<112> $P(\cdot)$ 는 예측 신호대잡음비가 음수인 경우 MMSE 추정값(\hat{A}_k)이 항상 양수임을 보장하기 위하여 사용된다. 상기 수학적식 26에 정의된 예측 SNR이 항상 양수는 아니기 때문에 상기 수학적식 31에 정의된 $P(\cdot)$ 를 사용하면 프라이오리 SNR가 항상 양수가 됨을 보장한다. 그 결과,

예측 SNR의 급격한 변화에도 프라이오리 SNR이 부드럽게 변화하도록 할 수 있다. 즉, 예측 SNR이 급격히 변화하는 경우 프라이오리 SNR은 1 프레임 지연되어 예측 SNR을 추종함으로써, 뮤지컬 잡음효과(musical noise effect)를 줄일 수 있다. 이에 대해서는 O. Capp에 의한 "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor, " *IEEE Trans. Speech and Audio Processing*, vol. 2, pp. 345-349, 1994』에 자세히 기재되어 있다.

<113> 57 단계에서는 상기 54 단계에서 얻어진 각 주파수대역별 진폭추정값을 이용하여 다음 수학적식 32에서와 같이 각 주파수대역별로 추정된 스펙트럼벡터를 생성한다.

<114> 【수학적식 32】 $\hat{S}_k(m) = \hat{S}_k^{(j)}(m)$

<115> 여기서, $\hat{\alpha}_k = \alpha_k$ 이다.

<116> 다시 도 1로 돌아가서, 목적원별 신호분리부(17)에서는 목적원 식별부(13)로부터 제공되는 m번째 프레임에서 목적원별 레이블벡터와 신호 추정부(15)로부터 제공되는 m번째 프레임의 각 주파수대역별로 추정된 스펙트럼벡터를 승산한 후, 역푸리에 변환을 수행하여 시간영역의 목적원 신호들로 분리한다. 이에 대하여 도 6을 참조하여 좀 더 세부적으로 설명하면 다음과 같다.

<117> 도 6에 있어서, 61 단계에서는 목적원 식별부(13)로부터 제공되는 클러스터의 수와 각 클러스터에 소속된 주파수대역의 인덱스에 따라 '1' 또는 '0'으로 표현되는 레이블벡터와 신호 추정부(15)로부터 제공되는 m번째 프레임의 각 주파수대역별로 추정된 스펙트럼벡터를 입력받는다.

- <118> 63 단계에서는 상기 61 단계에서 입력된 목적원별 레이블벡터와 추정된 스펙트럼벡터를 각각 승산하여 목적원별로 분리된 신호를 얻는다.
- <119> 65 단계에서는 상기 63 단계에서 얻어진 목적원별로 분리된 신호에 대하여 역푸리에변환을 수행하여 시간영역의 신호로 변환한다.
- <120> 신호 추정부(15)로부터 제공되는 m번째 프레임의 각 주파수대역별로 추정된 스펙트럼벡터를 이용한 목적원별 신호 분리는 k와 $\hat{S}_k(m)$ 의 적절한 페어링에 의해 수행되어진다. 상기 수학식 4에 기초하여 상기 수학식 32는 다음 수학식 33과 같이 표현될 수 있다.
- <121> **【수학식 33】** $\hat{S}_k(m) = \hat{S}_k^{(1)}(m) + \dots + \hat{S}_k^{(N)}(m)$
- <122> 또한, w-디스조인트 직교성 가정에 근거하여 상기 수학식 33은 다음 수학식 34로 간략화되어진다.
- <123> **【수학식 34】** $\hat{S}_k(m) = \hat{S}_k^{(j)}(m)$
- <124> 목적원별로 신호를 분리하기 위하여 각 목적원에 대응하는 레이블 벡터를 생성하고, 분리된 목적원별 신호 스펙트럼을 생성하기 위하여 목적원별 레이블 벡터를 추정된 스펙트럼벡터($\hat{S}(m)$)에 적용시킨다. j번째 목적원과 m 번째 프레임에서 k 번째 주파수대역에 존재하는 j번째 목적원 신호성분($S_k^{(j)}(m)$)을 매핑시키는 레이블 벡터($L_j(m)$)의 각 성분은 0 또는 1의 값을 가진다. 이에 따라 m 번째 프레임에서 j 번째 목적원에 대한 스펙트럼 벡터($\hat{S}_j(m)$)를 다음 수학식 35에서와 같이 얻을 수 있다.
- <125> **【수학식 35】** $\hat{S}_j(m) = L_j(m) \odot \hat{S}(m)$
- <126> 상기 수학식 35에 있어서 \odot 은 각 벡터에서 각 성분의 곱을 나타낸다.

- <127> 도 7은 본 발명에 따른 목적원 분리장치의 제2 실시예의 구성을 보여주는 블록도로서, 신호부재확률 계산부(71), 신호추정부(73), 목적원 식별부(75) 및 목적원별 신호분리부(77)로 이루어진다.
- <128> 도 7을 참조하면, 신호부재확률 계산부(71)에서는 제1 및 제2 마이크로폰 수신신호($X_1(m)$, $X_2(m)$)에 대하여 프레임별 m 번째 프레임에 대한 광역 신호부재확률($p(H_0 | X_1(m))$ 및 $p(H_0 | X_2(m))$)와 m 번째 프레임의 각 주파수빈에 대한 국지 신호부재확률($p(H_{0,k} | X_{1,k}(m))$ 및 $p(H_{0,k} | X_{2,k}(m))$)을 계산한다.
- <129> 신호추정부(73)에서는 제1 및 제2 마이크로폰 수신신호에 대하여 각각 신호부재확률 계산부(71)에서 계산된 국지 신호부재확률을 이용하여 프레임단위로 각 주파수대역별로 잡음신호가 제거된 스펙트럼 벡터를 추정하고, 잡음신호가 제거된 제1 및 제2 마이크로폰 수신신호($X_1(m)$, $X_2(m)$)를 생성한다.
- <130> 목적원 식별부(75)는 신호추정부(73)에서 잡음신호가 제거된 제1 및 제2 마이크로폰 수신신호($X_1(m)$, $X_2(m)$)를 이용하여 전술한 제1 실시예에서와 같이 클러스터링을 수행하여 목적원 및 각 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성한다.
- <131> 목적원별 신호 분리부(77)는 목적원 식별부(75)로부터 제공되는 m 번째 프레임에서 목적원별 레이블벡터와 신호 추정부(73)로부터 제공되는 m 번째 프레임의 각 주파수대역별로 추정된 스펙트럼벡터를 승산한 후, 역푸리에변환을 수행하여 시간영역의 목적원 신호들로 분리한다.
- <132> 즉, 상기한 바와 같은 구성은 제1 및 제2 마이크로폰 수신신호($X_1(m)$, $X_2(m)$)의 신호대잡음비가 큰 경우 적용할 수 있다.



<133> 다음, 본 발명에 따른 신호원 분리 알고리즘의 성능을 평가하기 위하여 음성신호에 대하여 실험을 행하였다. 미리 클린 음성 데이터를 에코가 없는 방에서 16 kHz 로 기록하고, 목적 신호를 생성하기 위하여 4명의 남성 화자가 발성한 컴퓨터 명령어와 한국어 이름에 대한 40 워드를 사용하였다. 각 화자가 발음한 40 워드를 4개 워드로 구성되는 10 개 파일로 분할하여 10 세트의 목적원을 준비하고, 백색 가우시안 잡음(White Gaussian Noise)을 광대역 잡음으로 사용하였다. 이때 신호대잡음비(SNR)에 따라서, ITU-T 권고안 P.830에 기재된 방법에 의해 신호 레벨을 조정한 다음 WGN을 샘플별로 부가하였다. 실제로 신호 혼합은 랩탑 컴퓨터에서 행해졌고, 신호원 식별 및 신호 추정은 프레임별로 수행되었다.

<134> 도 8 및 도 9a 내지 9c는 본 발명에 따른 신호원 분리 알고리즘의 실험 결과를 설명하기 위한 도면으로서, 도 8에서와 같이 제1 및 제2 마이크로폰(Mic1, Mic2)과 제1 내지 제4 목적원(Source1~Source4)이 사용된 경우, 도 9a는 제1 내지 제4 목적원신호(s1~s4), 도 9b는 제1 및 제2 마이크로폰 수신신호(x1, x2), 도 9c는 본 발명에 따른 목적원 분리 알고리즘에 의해 목적원별로 분리된 신호(u1~u4)를 각각 나타낸다. 도면을 참조하면, 본 발명에 따른 목적원 분리 알고리즘을 사용할 경우 잡음을 실시간으로 현저하게 감소시킬 수 있음과 동시에 목적원별로 신호를 보다 정확하게 분리할 수 있음이 입증되었다.

<135> 상기한 본 발명의 실시예는 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 한편, 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플라피디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되



어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.

<136> 또한, 본 발명에 따른 목적원 분리방법 및 장치는 마이크로폰 어레이가 부착된 자율주행 이동로봇 뿐만 아니라 사용되는 마이크로폰의 수가 적은 PDA(Personal Digital Assistant), 웹패드(WebPad) 또는 차량에 설치되는 휴대폰 등과 같이 사용자의 입과 단말기가 상당히 떨어진 환경에서도 적용가능하며, 이 경우에도 음성인식기의 성능을 대폭 향상시킬 수 있다. 또한, 고품질의 음성을 필요로 하는 모든 음성입력시스템에 적용가능하며, 동일한 환경에서 복수개의 음원이 존재하는 경우 각 음원을 분리하는 음원분리시스템에 사용될 수 있다. 또한, 통신 및 레이더 분야 등 기존의 센서 어레이가 사용되는 모든 분야에서 사용될 수 있다.

【발명의 효과】

<137> 상술한 바와 같이 본 발명에 따르면, 단지 두 개의 마이크로폰으로부터 수신되는 혼합신호로부터 실시간으로 잡음을 제거함과 동시에 임의의 N개의 목적원별 신호를 분리하는 것이 가능한 이점이 있다. 또한, 감쇠 및 지연시간을 고려하여 각 주파수대역에 대한 신호부재확률을 산출함으로써 w-디스조인트 직교성 가정을 위반하는 백색 가우시안 잡음이 혼합된 경우에도 적용할 수 있으므로 복원 음질의 향상을 도모할 수 있다. 또한, 주변 배경잡음으로 인하여 들리지 않는 목적원신호에 대해서도 신호부재확률을 이용하여 추정하는 것이 가능하다.

<138> 이상 도면과 명세서에서 최적 실시예들이 개시되었다. 여기서 특정한 용어들이 사용되었으나, 이는 단지 본 발명을 설명하기 위한 목적에서 사용된 것이지 의미 한정이나 특허청구범위에 기재된 본 발명의 범위를 제한하기 위하여 사용된 것은 아니다. 그러므로 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이



1020030020059

출력 일자: 2003/10/21

해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의해 정해져야 할 것이다.

【특허청구범위】

【청구항 1】

두 개의 센서로부터 수신되어 주파수영역으로 변환된 제1 및 제2 센서수신신호로부터 적어도 두 개 이상의 목적원별로 신호를 분리하는 방법에 있어서,

(a) 상기 제1 및 제2 센서수신신호 중 적어도 하나에 대하여 프레임별로 광역 신호부재 확률과 해당 프레임의 각 주파수대역에 대한 국지 신호부재확률을 계산하는 단계;

(b) 상기 (a) 단계에서 계산된 광역 신호부재확률에 의해 각 주파수대역별로 잡음신호를 제거한 스펙트럼벡터를 추정하는 단계;

(c) 상기 (a) 단계에서 계산된 국지 신호부재확률에 의해 신호가 존재한다고 판단되는 주파수대역의 신호값에 대하여 생성되는 감쇠와 지연시간을 이용하여 각 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 단계; 및

(d) 상기 (b) 단계에서 얻어진 각 주파수대역에 대하여 추정된 스펙트럼벡터와 상기 (c) 단계에서 얻어진 목적원별 레이블벡터를 각각 승산하여 목적원별로 분리된 신호를 획득하는 단계를 포함하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 2】

제1 항에 있어서, 상기 (a) 단계는

(a1) 소정 프레임의 각 주파수대역에 대한 가능비를 생성하는 단계;

(a2) 상기 (a1) 단계에서 생성되는 가능비를 각각 어프라이어리 확률과 승산하는 단계;

(a3) 상기 (a2) 단계에서 얻어지는 각 승산결과와 소정 값을 가산하는 단계;

(a4) 상기 (a3) 단계에서 얻어지는 각 가산결과의 역수를 취하여 각 주파수대역의 국지 신호부재확률로 설정하는 단계; 및

(a5) 상기 (a4) 단계에서 얻어지는 각 국지 신호부재확률을 승산하여 상기 소정 프레임의 광역 신호부재확률로 설정하는 단계를 포함하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 3】

제1 항에 있어서, 상기 (b) 단계는

(b1) 상기 (a) 단계에서 계산된 광역 신호부재확률에 의해 잡음신호만 존재하는 프레임에 대해서는 각 주파수대역별로 잡음신호전력을 갱신하는 단계;

(b2) 상기 (a) 단계에서 계산된 광역 신호부재확률에 의해 목적원신호와 잡음신호가 공존하는 프레임 또는 상기 (b1) 단계에서 잡음신호전력이 갱신된 프레임에 대하여 각 주파수대역의 스펙트럼성분의 진폭을 추정하는 단계; 및

(b3) 상기 (b2) 단계에서 추정된 각 주파수대역의 스펙트럼성분의 진폭을 이용하여 각 주파수대역별로 목적원신호전력을 갱신하는 단계를 포함하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 4】

제3 항에 있어서, 상기 각 주파수대역의 스펙트럼성분의 진폭은 최소평균자승오차 추정 알고리즘에 의해 추정되는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 5】

제1 항에 있어서, 상기 (c) 단계는

(c1) 상기 (a) 단계에서 계산된 각 주파수대역별 국지 신호부재확률을 제1 문턱치와 비교하여 신호가 존재하는 주파수대역을 식별하는 단계;

(c2) 상기 (c1) 단계에 의해 신호가 존재한다고 판단되는 주파수대역에 대하여 상기 제1 및 제2 센서수신신호의 비로부터 감쇠와 지연시간을 생성하는 단계;

(c3) 상기 (c2) 단계에서 생성되는 감쇠와 지연시간으로 이루어지는 혼합파라미터를 클러스터링하는 단계; 및

(c4) 상기 (c3) 단계에서의 클러스터링한 결과 생성되는 각 클러스터에 소속된 혼합 파라미터를 이용하여 목적원과 해당 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 단계를 포함하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호 분리방법.

【청구항 6】

제5 항에 있어서, 상기 혼합 파라미터의 클러스터링은 소프트 K-평균 알고리즘에 의해 수행되는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 7】

제5 항에 있어서, 상기 혼합 파라미터의 클러스터링은 하드 K-평균 알고리즘에 의해 수행되는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 8】

제5 항에 있어서, 상기 (c1) 단계는 상기 주파수영역으로 변환된 제1 및 제2 센서수신신호 중 상기 (a) 단계에서 신호부재확률이 계산된 신호에 대하여 수행되는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 9】

제5 항에 있어서, 상기 (c1) 단계는 상기 (b) 단계에서 잡음신호가 제거된 제1 및 제2 센서수신신호에 대하여 수행되는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리방법.

【청구항 10】

제1 항 내지 제9 항에 기재된 방법을 수행할 수 있는 프로그램을 기재한 컴퓨터로 읽을 수 있는 기록매체.

【청구항 11】

두 개의 센서로부터 수신되어 주파수영역으로 변환된 제1 및 제2 센서수신신호로부터 적어도 두 개 이상의 목적원별로 신호를 분리하는 장치에 있어서,

상기 제1 및 제2 센서수신신호 중 적어도 하나에 대하여 프레임별로 광역 신호부재확률과 해당 프레임의 각 주파수대역에 대한 국지 신호부재확률을 계산하는 신호부재확률 계산부;

상기 신호부재확률 계산부에서 계산된 광역 신호부재확률에 의해 각 주파수대역별로 잡음신호를 제거한 스펙트럼벡터를 추정하는 신호추정부;

상기 신호부재확률 계산부에서 계산된 국지 신호부재확률에 의해 신호가 존재한다고 판단되는 주파수대역의 신호값에 대하여 생성되는 감쇠와 지연시간을 이용하여 각 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 목적원 식별부; 및

상기 목적원 식별부에서 얻어진 목적원별 레이블벡터와 상기 신호추정부에서 얻어진 각 주파수대역에 대하여 추정된 스펙트럼벡터를 각각 승산하여 목적원별로 분리된 신호를 목적원별 신호분리부를 포함하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리장치.

【청구항 12】

제11 항에 있어서, 상기 목적원 식별부는 상기 감쇠와 지연시간으로 이루어지는 각 주파수대역별 혼합 파라미터를 소프트 K-평균 알고리즘을 이용하여 클러스터링한 후, 클러스터링한 결과 생성되는 각 클러스터에 소속된 혼합 파라미터를 이용하여 목적원과 해당 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리장치.

【청구항 13】

제11 항에 있어서, 상기 목적원 식별부는 상기 감쇠와 지연시간으로 이루어지는 각 주파수대역별 혼합 파라미터를 하드 K-평균 알고리즘을 이용하여 클러스터링한 후, 클러스터링한 결과 생성되는 각 클러스터에 소속된 혼합 파라미터를 이용하여 목적원과 해당 목적원에 속하는 주파수대역으로 이루어지는 목적원별 레이블벡터를 생성하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리장치.

【청구항 14】

제11 항에 있어서, 상기 신호추정부는 각 주파수대역의 스펙트럼에 대한 진폭을 최소평균자승오차 추정 알고리즘에 의해 추정하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리장치.

【청구항 15】

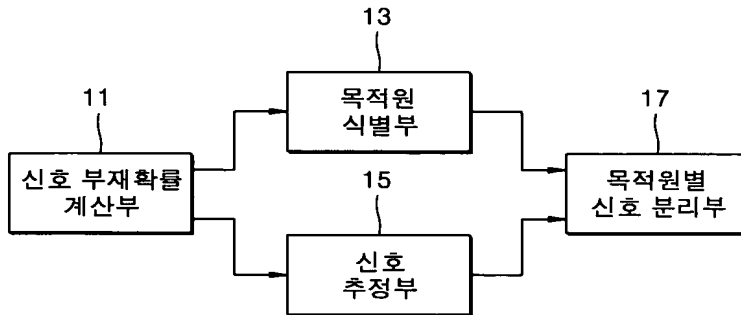
제14 항에 있어서, 상기 신호추정부는 상기 광역 신호부재확률을 이용하여 잡음신호전력이 갱신된 잡음신호만 존재하는 프레임 및 잡음신호와 목적원신호가 공존하는 프레임에 대해서 목적원신호전력을 매 프레임마다 갱신하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리장치.

【청구항 16】

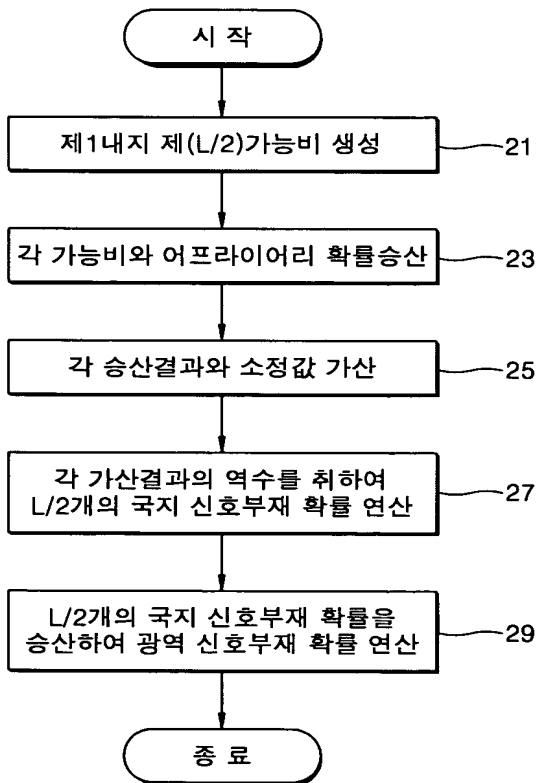
제11 항에 있어서, 상기 제1 및 제2 센서수신신호의 신호대잡음비가 큰 경우에는 상기 목적원 식별부에서는 상기 신호추정부에 의해 잡음이 제거된 상기 제1 및 제2 센서수신신호를 이용하여 크기 감쇠 및 지연시간을 계산하는 것을 특징으로 하는 두 개의 센서를 이용한 목적원별 신호분리장치.

【도면】

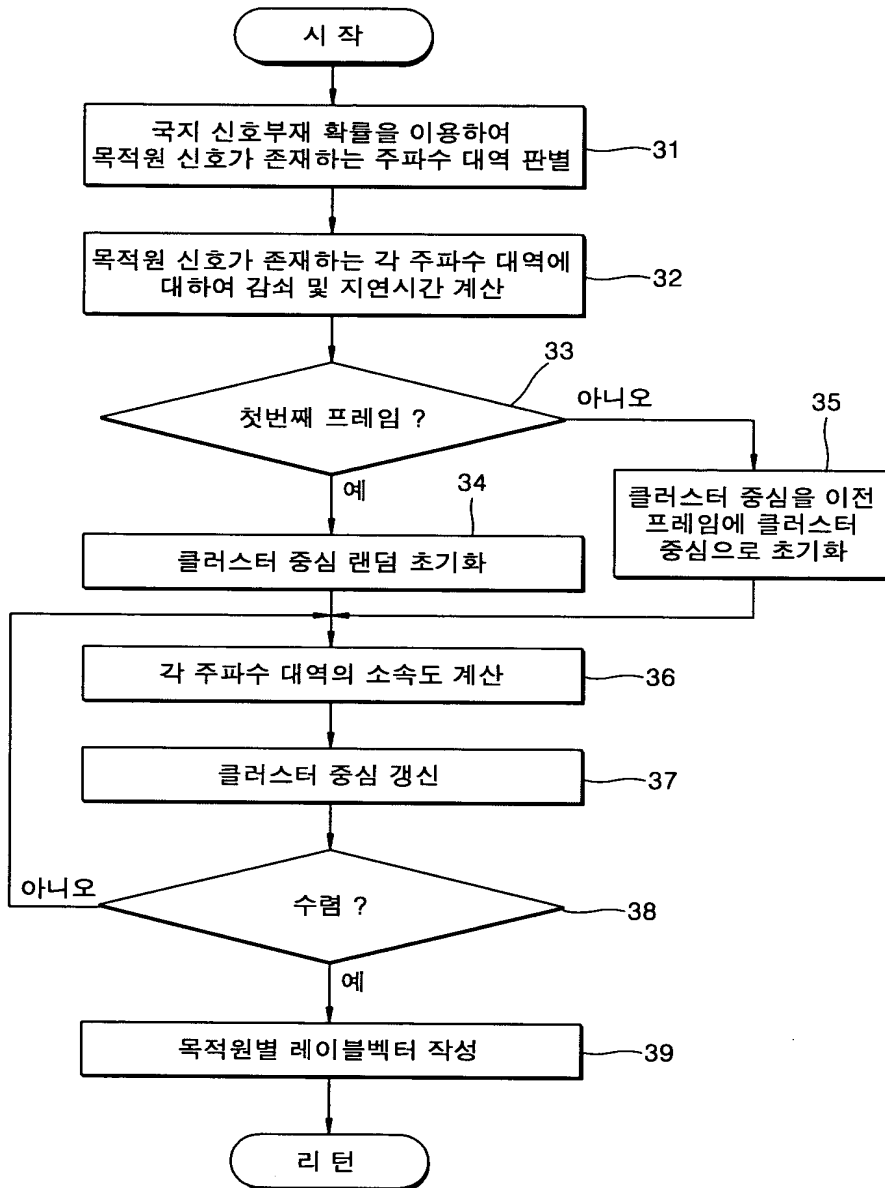
【도 1】



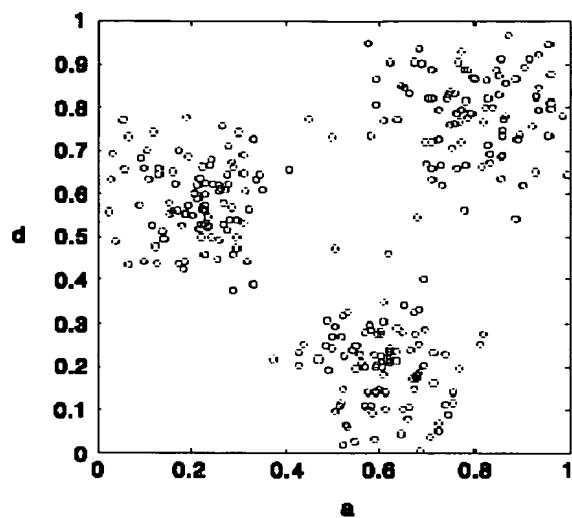
【도 2】



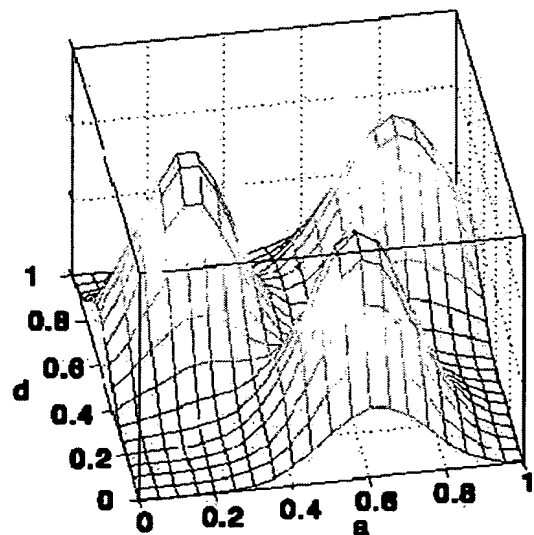
【도 3】



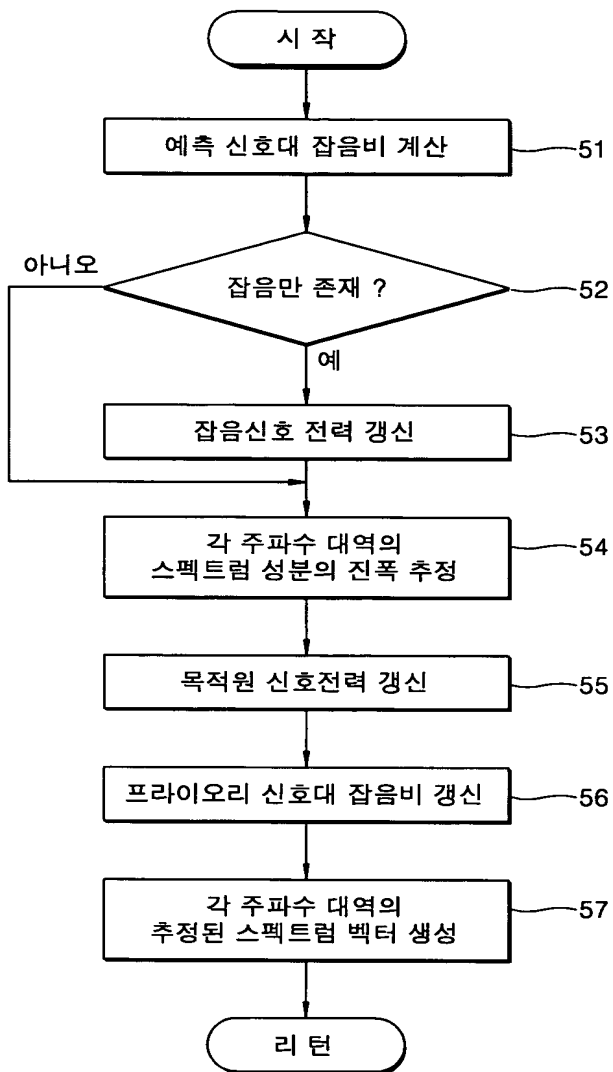
【도 4a】



【도 4b】

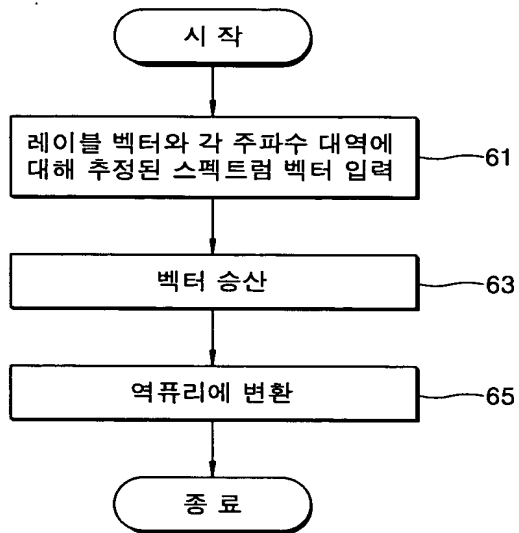


【도 5】

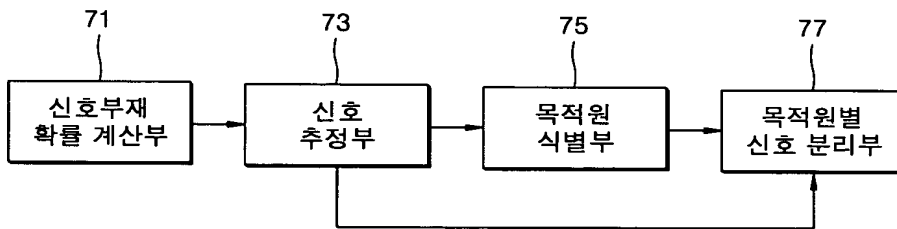




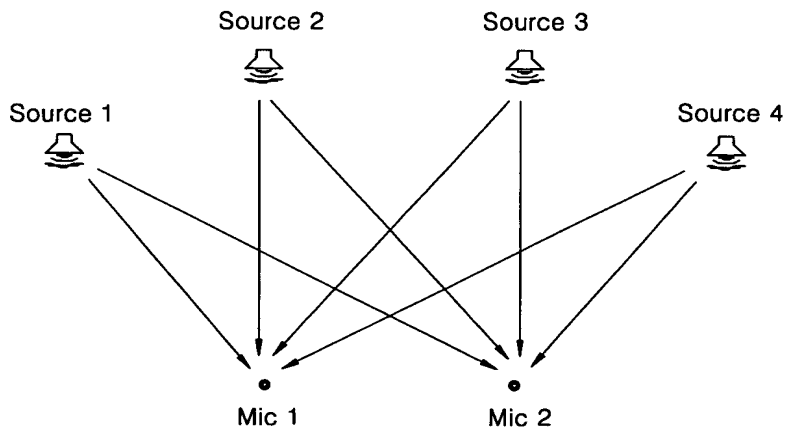
【도 6】



【도 7】

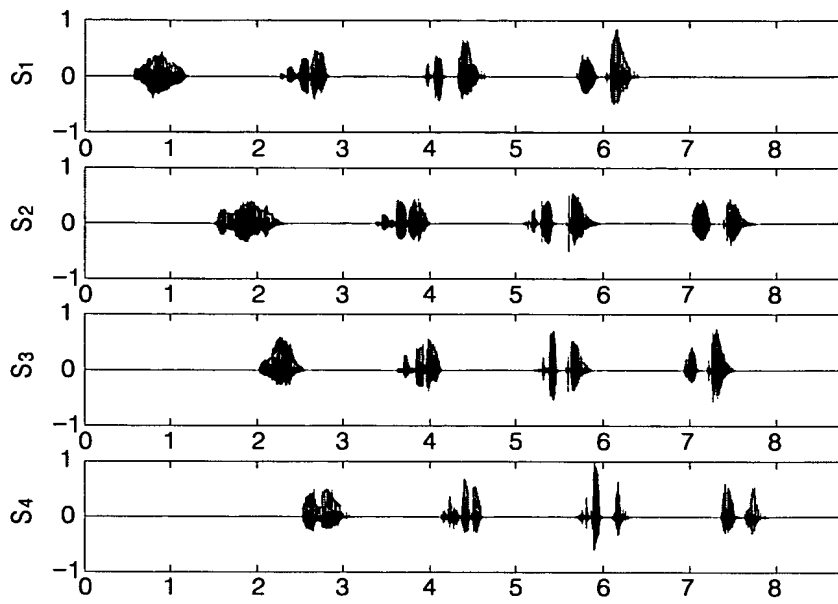


【도 8】

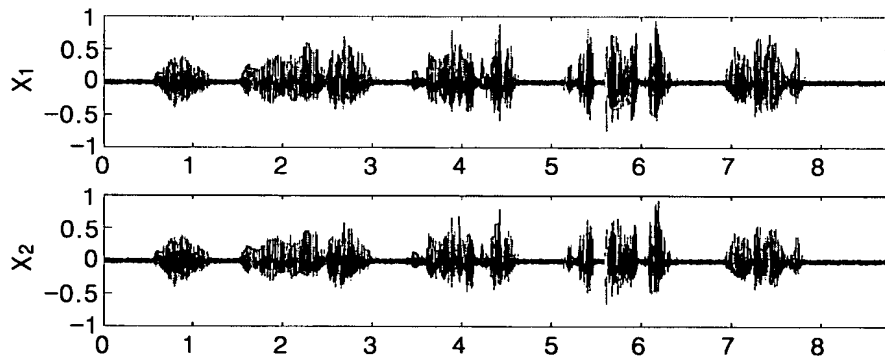




【도 9a】



【도 9b】



【도 9c】

